DOCUMENT RESUME

ED 374 802 IR 055 123

AUTHOR Conway, Paul; Weaver, Shari
TITLE The Setup Phase of Project Open Book: A Report to the Commission on Preservation and Access on the Status of an Effort To Convert Microfilm to Digital Imagery.
INSTITUTION Commission on Preservation and Access, Washington, DC.
PUB DATE Jun 94
NOTE 29p.
AVAILABLE FROM Commission on Preservation and Access, 1400 16th Street, N.W., Suite 740, Washington, DC 20036-2217 ($10).
PUB TYPE Reports - Descriptive (141)

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS *Access to Information; Administration; College Libraries; Databases; Higher Education; *Indexing; *Information Storage; Library Materials; Microfilm; Models; *Preservation; *Production Techniques; Program Implementation
IDENTIFIERS *Digital Imagery; *Library Materials Conservation; Yale University CT

ABSTRACT

Digital image quality, indexing structures, and production workflow were the three central issues examined during the second phase--the set-up phase--of Project Open Book, a major effort by Yale University Library to explore the usefulness of digital technologies for preserving and improving access to deteriorating documents. This report outlines the incremental progress made during the set-up phase and describes the results of Yale's investigation into the administrative requirements for high-volume production conversion. In 1993, Yale set up and evaluated components of an in-house production-conversion facility, converted and indexed 100 volumes in a test run, and prepared for the conversion from preservation microfilm of the next 3,000 volumes of a projected 10,000-volume digital library. A production workflow plan was developed that stressed the importance of working with vendors, training staff, and broad administrative support. The set-up phase focused intensely on process quality, efficient ways to use existing database technology, and specific steps in the conversion process. The production-conversion phase will define requirements for conversion and validate the conversion model. (Contains 11 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 374 802

The Commission on Preservation and Access

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

THE SETUP PHASE OF PROJECT OPEN BOOK

A Report to the Commission on Preservation and Access
On the status of an effort to convert microfilm to digital imagery

by

Paul Conway
Head, Preservation Department

and

Shari Weaver
Assistant to the Director
Library and Administrative Systems
Yale University Library

IROSS 133

June 1994

"PERMISSION TO REPRODUCE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY

M.K. Sitts

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

1400 16th Street, N.W., Suite 740, Washington, D.C. 20036-2217 • (202) 939-3400

A private, nonprofit organization acting on behalf of the nation's libraries, archives, and universities to develop and encourage collaborative strategies for preserving and providing access to the accumulated human record.

The Setup Phase of Project Open Book

A Report to the Commission on Preservation and Access

On the status of an effort to convert microfilm
to digital imagery

by

Paul Conway
Head, Preservation Department
and
Shari Weaver
Assistant to the Director
Library and Administrative Systems
Yale University Library

June 1994

Published by
The Commission on Preservation and Access
1400 16th Street, NW, Suite 740
Washington, DC 20036-2217

June 1994

Reports issued by the Commission on Preservation and Access are intended to stimulate thought and discussion. They do not necessarily reflect the views of Commission members.

Additional copies are available from the above address for $10.00. Orders must be prepaid, with checks made payable to "The Commission on Preservation and Access," with payment in U.S. funds.

This paper has been submitted to the ERIC Clearinghouse on Information Resources.

The paper in this publication meets the minimum requirements of the American National Standard for Information Sciences-Permanence of Paper for Printed Library Materials ANSI Z39.48-1992.

COPYRIGHT 1994 by the Commission on Preservation and Access. No part of this publication may be reproduced or transcribed in any form without permission of the publisher. Requests for reproduction for noncommercial purposes, including educational advancement, private study, or research will be granted. Full credit must be given to the author(s) and The Commission on Preservation and Access.

Executive Summary

Digital image quality, indexing structures, and production workflow were the three central issues examined during the second phase--the setup phase--of Project Open Book, a major effort by the Yale University Library to explore the usefulness of digital technologies for preserving and improving access to deteriorating documents. This report to the Commission on Preservation and Access outlines the incremental progress made during the setup phase, conducted under contract to the Commission, and describes the results of Yale's investigation into the administrative requirements for high-volume production-conversion. In past years, the Commission has supported the development of the master plan for Project Open Book and the first phase of the project.

In 1993 Yale set up and evaluated the components of an in-house production conversion facility, converted and indexed 100 volumes in a test run, and prepared for the conversion from preservation microfilm of the next 3,000 volumes of a projected 10,000 volume digital library. The setup phase built upon the first phase of the project--the organizational phase--in which Yale conducted a formal bid process and selected Xerox Corporation to serve as its principal partner in Project Open Book.

During the setup phase Yale acquired a single integrated conversion workstation through Xerox, including microfilm scanning hardware and associated conversion and enhancement software, tested and evaluated this workstation, and made the transition to a fully engineered production system. In the process of developing a production workflow plan for Project Open Book, the project team defined three conditions that must be present in the library in order for the project to proceed, including a close working relationship with technology vendors, the recruitment and training of skilled staff, and broad administrative support in the library and the university.

The setup phase focused intensely on three related issues. The project team first defined the quality of the conversion process from microfilm by understanding the interaction and impact of three groups of factors on image quality: the characteristics of the original source materials reformatted on microfilm, the characteristics of the microfilm itself, and the technical limitations of the conversion equipment. Second, the project team identified efficient ways to use existing database technology to develop structural indexes for large, complex, and varied published source materials stored on microfilm. Third, the project team outlined in some detail the specific steps of the conversion process, estimated the productivity rates of each of these steps, and identified the major factors affecting the rate of conversion. The setup phase also resulted in a plan to address two issues that are critical to the overall success of Project Open Book; namely, a criteria-driven selection process and an approach to evaluating the findings of the project.

In the third phase of the project--the production-conversion phase--Yale will define the requirements for effective and economical digital conversion from microfilm images, validate the production conversion model developed during the setup phase, and seek to optimize the efficiency of the digital conversion process.

The Setup Phase of Project Open Book

Introduction

During 1993, Yale University Library set up and evaluated the components of an in-house production conversion facility, converted and indexed 100 volumes in a test run, and is now poised to begin the conversion from preservation microfilm of the next 3,000 volumes of a projected 10,000 volume digital library. In the first phase of the project--the organizational phase--Yale conducted a formal bid process and selected Xerox Corporation to serve as its principal partner in Project Open Book. During the second phase of the project--the setup phase--Yale acquired a single integrated conversion workstation through Xerox, including microfilm scanning hardware and associated conversion and enhancement software; tested and evaluated this workstation using prototype production-conversion and indexing software developed by Xerox; and made the transition to a fully engineered production system. The Commission on Preservation and Access provided support for the planning of Project Open Book as well as for the organizational and setup phases of the project.

The setup phase involved the in-depth investigation of quality assurance and production workflow issues, the initial development of selection guidelines for the identification of preserved materials appropriate for conversion, and the formation of a preliminary plan for evaluating Project Open Book through significant input from Yale's research community. The setup phase also laid the administrative groundwork for full-scale production conversion which will take place in stages three and four. With partial funding from the National Endowment for the Humanities, the Yale Library is now prepared to select material for conversion from microfilm to digital imagery, to install and manage a multi-platform networked conversion system, and to implement a rich evaluation process over the remaining life of Project Open Book.

This report will outline the steps undertaken in the setup phase, describe the results of our investigation into issues of project administration, image quality, and the management of high-volume production-conversion, and summarize the project's approach to selection and evaluation.

Background

The June 1991 report entitled *From Microfilm to Digital Imagery* is the master plan for Project Open Book, a major effort in the Yale University Library to explore the usefulness of digital technologies for preserving and improving access to deteriorating documents.[1] The project is founded on a vision of the research library as an institution whose mission is to generate, preserve, and improve for its clients ready access--both intellectual and physical--to recorded knowledge. The place of electronic tools and information sources in the library of the future will depend on how well (or how poorly) they measure up against this mission. Digital image documents are among the various types of information that

most likely will find a place in the access-oriented library committed to making available sources in electronic form. Yale's long-standing leadership in the preservation of paper-based library materials drives us to test the feasibility of using digital imaging as a preservation tool.

The purpose and scope of Project Open Book are designed to lead research libraries like Yale closer to a more general "ideal" model of the role of digital image documents in the transforming library, originally outlined in *From Microfilm to Digital Imagery*. The model posits the existence of an image document library that is created from multiple sources and with multiple uses. We imagine libraries will generate digital image documents from film and paper for preservation purposes as well as for more general reasons, such as the creation of reserve materials or customized textbooks. Libraries may also acquire image documents from external sources, such as service bureaus hired to reformat preservation materials or directly from publishers or vendors. After digital conversion, libraries may opt to move the film and paper to remote storage. Users may then print documents on demand from the image library, browse them at a workstation, or reformat them, say, by generating microfilm or by submitting them to character recognition processes. The model posits that the quality of the image documents that libraries maintain will depend in large part on the expected mix of these various uses in both the long- and the short-term.

In Project Open Book, Yale will convert 10,000 volumes into digital image form. This number is an order of magnitude larger than the number of documents originally converted in Cornell University's pioneering CLASS Project and our intention is to explore the effects of scale on emerging preservation imaging systems.[2] In addition, rather than scanning directly from the original paper document, the purpose of the Yale project is to convert material from frames of 35 millimeter microfilm to digital images and thus explore the promise that once we have preserved materials on film we can eventually and satisfactorily convert those documents into digital form. Several other working hypotheses also help to define and limit the scope of Project Open Book, including:

- Microfilm is satisfactory as a long-term medium for preserving content, even if it falls short as an access medium;
- Digital imagery can improve access to recorded knowledge though printing and network distribution at a modest incremental cost over microfilm;
- Researchers will demand greater access to digital image libraries if they contain thematically-related materials than if image documents are randomly dispersed topically;
- Capturing and storing documents in digital image form is a necessary step leading to even further improvements in access (e.g., through the application of optical character recognition).

Taken together and verified in efforts like Project Open Book, these various hypotheses may lead to the specific conclusion that research libraries will choose, given a mix of flexible technology, to maintain information on microfilm for long-term preservation and in digital image form for ease of access. Regardless of its findings on the function of preservation microfilm, Project Open Book promises to identify the institutional preconditions for the full integration of digital imaging technology into library preservation processes.

Yale completed the first phase of the project--the organizational phase--in 1992. As part of this phase, Yale established a Steering Committee, including several faculty members, and created a project team. The bulk of the organizational phase was devoted to a formal bid process which led to the selection of the Xerox Corporation as a principal partner and the identification of areas of risk and uncertainty and key issues to be addressed over the life of the project. One benefit of the selection process was the involvement of a large number of Yale staff both from the library and computer center in the process of working with the competing vendors to develop their bids. These staff contributed substantially to the analysis of the project requirements and gained knowledge and expertise in imaging systems. A detailed summary of the organizational phase is contained in the report, *The Organizational Phase of Project Open Book*.[3]

The first phase confirmed that the management of complex documents in image form is a general problem crying for solution in many areas. It is not confined to library preservation, to libraries, or even to academic institutions. Although the market for imaging products is thus potentially broad, our experience suggests that it is still maturing. Development of the market will depend on many factors, but in our view the adoption of digital imaging technology by libraries will depend on the successful resolution of several key issues, in particular the quality of digital conversion from microfilm, the cost-effective production of image data, and the indexing of image files to facilitate browsing. These three issues, along with the continuing need to identify and set selection criteria formed the heart of the setup phase, just concluded. Yale's interactions with Xerox during the organizational phase generally validated the original master plan for Project Open Book.

The Setup Phase

The overall goals of the second phase of Project Open Book--the setup phase--were to assemble, test, and evaluate the basic operating elements of the imaging architecture for the project. Installed components include a single conversion station built around an Intel 486/33 platform, running MS-DOS® and MS Windows 3.1® and a pre-release version of the digital library management software which was upgraded during the course of the setup phase to the commercially available product, Xerox Documents on Demand. In addition to the PC outfitted with a 19-inch high-resolution monochrome monitor, the conversion workstation includes a Mekel M400 microfilm scanner, a local device for

writing digital data to 5.25 inch magneto-optical (re-writable) disks, a Hewlett Packard LaserJet local printer for proofing newly scanned images, a Xerox 4030 draft printer for hardcopy output from the Xerox software, network access to a high-volume, high-resolution Xerox DocuTech printer, and, for bibliographic information about the documents in image form, access to Orbis, the Yale University Library online catalog. An appendix outlines the components of the single-workstation system installed at Yale during the setup phase.

The Mekel M400 scanner is controlled by Amitech Corporation's TurboScan® software program. It scans, compresses, and saves sequentially numbered image files, at a maximum rate of one frame every three seconds. A scanner options menu activates a number of image processing functions, including automated document skew correction, scaling, cropping and rotation. A date and time stamp are automatically assigned to each image file by DOS. File names are controlled by the technician and a batch information file keeps track of the number, type, and location of associated files. Typically each batch is one volume.

A digitally scanned raster image is essentially an electronic "photograph" of a document, divided into a "grid" composed of thousands of minuscule picture elements or pixels, each one representing a very small area on the original document. Unlike alphanumeric data, raster images consist of binary 1s and 0s that in themselves carry no intelligence, and therefore cannot be queried in terms of what information the image represents. At the point of scanning, each pixel is registered with an average brightness called a grey level. In order to reduce the amount of space and computer memory required to store and process full grey-scale information, each pixel is stored as being either black or white, rather than as a grey level. Scanners typically use a method called thresholding to convert the grey pixel to a black or white pixel. The scanner compares the brightness of each pixel with a given threshold value. If the pixel is brighter than the threshold value, then the pixel is converted to white, otherwise black.

The Mekel scanner is equipped with a special peripheral called the Scan Optimizer® from Image Processing Technologies to improve the quality of this process with microfilm. The Scan Optimizer looks for transitions between dark areas and light areas to define their edges. This method works well with faded documents, poor quality originals on film, and other low-contrast images where it is difficult or impossible to define a specific threshold between black and white. The Scan Optimizer also permits the TurboScan software to define the beginning of a frame by recognizing the transition of the page from white (clear) to black (on negative microfilm).

A hand-held controller permits the technician to operate the Scan Optimizer interactively while scanning microfilm. The controller is used to set the sensitivity level (the amount of information captured), thickness (the relationship between lines and noise), threshold

levels (equivalent to fine tuning after the thickness level has been established), brightness levels (which determine the extent of black fill-in), and filters (which remove small, speckle noise and thin, broken lines). The combination of these settings determines the extent to which overall conversion quality is maximized, when quality is defined as a faithful reproduction of the original microfilm image. The Scan Optimizer allows for the establishment of up to ten "pre-sets" which store technician-defined combinations of settings that can then be applied to a particular batch of images by way of the hand-held controller.

In the course of converting 100 volumes from 35 millimeter preservation microfilm to digital image documents, Yale's project management team gained considerable expertise with the conversion equipment and the digital image software from Xerox and associated vendors. The team experimented widely with Scan Optimizer settings and identified a group of "pre-sets" that can be applied to microfilm images with varying characteristics. Additionally, the team identified the administrative "pre-conditions" necessary for successful implementation and achieved provisional definitions of the quality control, structural indexing, and workflow processes necessary for production-level conversion in subsequent phases of the projects. The team also established the broad parameters governing selection of materials for conversion that will lead to the definition of a specific selection process and honed its understanding of the evaluation criteria for the project as a whole. The following sections describe how Project Open Book has addressed each of these issues.

Pre-conditions for Project Implementation

Planning in an environment of change requires close coordination and flexibility. In the process of developing an implementation plan for Project Open Book, the project team had to come to terms with three conditions that must be present in the library in order for the project to proceed. A **close working relationship with vendor/manufacturers** of hardware and software components is one crucial pre-condition. Yale University's partnership arrangement with the Xerox Corporation, which was forged in a complex bidding process during the first phase of the project, provides a number of advantages for the library. Among these are direct access to technology laboratories and research scientists for problem solving, technical support and training on new equipment, advance notice of new and emerging product lines, and detailed information on corporate marketing strategies. Yale has benefitted from its partnership with Xerox by being able to project with greater certainty the directions in which the imaging industry is developing and to plan accordingly. The partnership arrangement with Xerox challenged Yale to balance its needs for the project with a corporation functioning within an increasingly competitive information technology industry. Library administrators exercising the partnership option must be prepared for complex negotiations and development cycles that are driven more by corporate bottom lines than library project management requirements.

Another requirement for project implementation is the presence of **skilled staff** who understand the characteristics of the source media to be converted to digital images and are capable of mastering the capabilities of the newly acquired technology. During the setup phase of the project, a project team of existing library staff donated time for administrative support or were detailed temporarily from core duties to undertake conversion tests. The availability of technical staff familiar with how preservation microfilm is created, processed, and used was essential to the successful completion of the phase. It became readily apparent, however, that near-term personnel needs could not be met with existing staff. Moreover, the existing technical job classification system at Yale has proven to be insufficient for the recruiting and training of qualified technicians to undertake the daily routine of digital imaging. In anticipation of a long-term commitment to digital image conversion, Yale is working to develop an entirely new job family of technicians and to hire project staff into this family at the appropriate skill levels needed for the project. Matters of personnel administration can be significant barriers to the implementation of imaging projects if not managed pro-actively.

Finally, successful project implementation depends upon **administrative support and leadership**. During the organizational phase of the project, the vendor selection process allowed Yale to involve a large number of staff both from the library and the computer center, and particularly during the development of the project's requirements, to cultivate their knowledge and expertise. In the setup phase of Project Open Book, the wide involvement of staff and faculty continued in the form of special demonstrations and briefings, consultation on technical processes, and wide ranging discussions on selection criteria and evaluation strategies. The project is now a central component of the library's vision of its emerging support for electronic research resources and is a continuing focus for fund raising and outreach to the library community. The burgeoning interest in the conversion of library materials to digital images is but one part of a larger transformation of library services. Relationships among publishers, libraries, and patrons are being fundamentally redefined in the face of information technology, and it is the responsibility of library administrators to manage this transformation.

Quality Digital Conversion

One of the central questions addressed through Project Open Book is: "What is the highest possible quality of microfilm conversion that can be accomplished within the limitations of existing technology?" Defining quality in the imaging environment is not a straightforward matter.[4] In the absence of computer-assisted image calibration that automatically sets filters and enhancement parameters to achieve maximum data conversion, the standard of acceptable quality in the imaging environment has been the judgment of experienced system operators. Even the effective use of calibration tools developed by the Association for Information and Image Management (AIIM), for example, ultimately depends on how any particular test target image appears on the

screen.[5]

Conversion from microfilm to digital imagery entails establishing procedures that mitigate the inevitable three-way interaction among the characteristics of the original source materials included on the microfilm, the technical characteristics of preservation microfilm itself, and the capabilities and limitations of the hardware/software conversion system. The setup phase of Project Open Book demonstrated that achieving quality conversion in a production-conversion environment, even when quality is rigorously defined, entails some measure of compromise between the capabilities of the conversion system and the demands of efficient production. It is essential that project administrators maintain thorough documentation, including logs of benchmarking runs and equipment tests, on the nature of the choices made to limit conversion quality in the interests of cost-effectiveness.

The setup phase of Project Open Book identified the major **characteristics of original source materials** that, when reformatted on preservation microfilm, complicate the production-conversion process. Library materials, be they books, journals, manuscripts, maps, or photographs, present significant challenges for digital image conversion. Unlike office documents, library materials vary greatly in size, contrast, color, and condition. Cornell University's CLASS project, as well as other ongoing pilot imaging projects, demonstrate the capability of imaging technology to achieve near-facsimile reproduction of paper-based materials. Obtaining the highest quality conversion of library materials in a production environment, however, requires that compromises be made in overall quality. The following are among the most significant characteristics of the original source document that affect conversion from microfilm.

- Tone and *physical condition* of paper. Books with yellowed and/or brittle edges, dog-eared pages, and uneven fore-edges tend to confuse the edge detection software equipped with the scanner. If the threshold level is set low enough to compensate for page edges that are not crisply defined, other discrepancies in the film, such as splices between frames, may cause the software to malfunction.

- The size, clarity, and contrast of the *typeface* are significant variables in the scanning process.

- The presence of *illustrations* such as line drawings, engravings, and half-tones may require special scanner settings which may impact the readability of accompanying text. In determining filter settings, a decision must be made about the intellectual significance of the illustrations in terms of the text and adjustments made accordingly.

- *Foldouts* such as maps and illustrations that must be reproduced across the entire 35 millimeter frame will inevitably slow digital conversion as settings must be adjusted

to accommodate full-frame imaging and then reset to resume the conversion of individual pages.

- Many *colors* do not reproduce well on high-contrast preservation microfilm. If color is integral to the understanding of the content, as is the case for maps, binary digital image conversion (black and white) most likely will not be appropriate.

The **characteristics of document images reproduced on preservation microfilm** have a tremendous impact on the quality of the resulting conversion as well as on the rate at which that conversion can take place. Preservation microfilm is, by nature, a high-contrast medium for reproduction. First generation film, used as the master negative and to create second generation copies, is panchromatic, extremely fine grain, silver-gelatin type document recording film. Second-generation film, used for the production of positive use copies and also used by Project Open Book for digital image conversion, is non-reversing (negative) silver-gelatin non-perforated polyester-based duplicating film. The RLG Preservation Microfilming Handbook specifies separate film stocks for materials with standard black on white text and for materials containing halftones or continuous tone illustrations.[6] The following are the most significant film characteristics.

- *Resolution* is the ability of the microfilming system (camera, lens, and film) to record fine detail. Resolution is measured on microfilm by viewing the reproduced test target containing test patterns arranged in groups of horizontal and vertical lines of specific size and spacing. The RLG handbook, which describes and interprets the relevant ANSI/AIIM standards, specifies that a good quality microfilm system should be able to register 120 line pairs per millimeter on duplicate negative film. The findings of the setup phase of Project Open Book reinforce the need to examine closely the methodology for interpreting resolution when micro-reproduced test targets are converted to digital image. AIIM has recommended the use of resolution indicators on the IEEE Std 167A-1987 facsimile test chart for calibrating digital image scanners.[7]

- *Density* refers to the opacity of the film. On master and duplicate negative film the maximum density--or background density--is the dark part of the image whereas the minimum density--or base plus fog--is the clear part of the film on which there is no image. The actual background density of the negative is in part determined by the contrast of the original materials. The RLG handbook specifies that the maximum density of film containing high contrast originals fall within the range of 1.0 to 1.3. Overall, microfilm should not have maximum densities below 0.80. Project Open Book has validated these density recommendations with the additional proviso that consistent average density readings within a single volume is a far more significant determinant of image conversion quality than any particular average density value within the range of 1.0 to 1.3.

- *Reduction ratio* is intimately related to the quality of conversion because the ratio determines the amount of visual data stored on the microfilm that is available for digital conversion. Library materials reproduced on microfilm are reduced to a specific ratio that may range from 8:1 to 14:1. The exact reduction ratio for a given volume is a crucial variable that the imaging system needs to reproduce faithfully the size of the original document. Although the RLG handbook permits libraries to film materials either at fixed or variable reduction ratios, variation of ratios within a volume will cause production delays.

Ultimately, the quality of the digital image conversion is determined by the specific **capabilities and limitations of the digital conversion hardware and software** as they interact with the characteristics of the original source materials reproduced on preservation microfilm. Among the most important of these capabilities in the Project Open Book conversion system are the following.

- Any scanning device has a finite limit on the *amount of data* that can be captured about any given surface area. By necessity, this limit determines the top range of image quality. The maximum capacity of the CCD array (charge couple device) of the Mekel scanner is 7042 pixels per inch (ppi) at 16 millimeters and 3694 ppi at the full 35 millimeter width. The conversion of microfilm at 600 dpi is actually a process of adjusting the capacity of the CCD array in relation to the reduction ratio of the materials preserved on film. Six-hundred dpi resolution, therefore, is actually a software-controlled mathematical artifact of the scanning process.

- The Mekel scanner and its associated software add-ons equip the technician with a variety of filters and threshold settings that determine the nature of the data conversion process. Inadequate *documentation* on obtaining high quality results, however, requires "trial and error" processing to understand the interaction of the filters and system settings.

- In its present configuration, the Mekel scanner converts each frame of microfilm with a uniform set of filter settings. The scanner does not have the capability of creating a combined digital image with one set of filters for text and another set of filters for "window" containing an illustration on the same physical page. The lack of a *"windowing" capability* means that the technician must make choices in the process of conversion about which type of image content (text or illustrations) to emphasize.

The ANSI/AIIM standard on monitoring microfilm scanner image quality recognizes the technical limitations of current conversion technology and identifies three principles of image quality control. The standard, designed expressly to support the conversion of

business documents, recommends establishing quality references defining "good output" from a digital image system, conducting scanner testing before and after converting each batch of microfilm, and maintaining thorough records of testing and calibration. Of these three, the concept of "quality reference" has proven to be the most challenging for the production team during the setup phase precisely because the existing standard presents a circular argument for identifying quality conversion. "If the **quality** of the digitized test target images is of **appropriate quality**, [emphasis added] ... this becomes the 'quality standard'."[8] The standard calls for scanning a suite of test targets, printing out the results on paper and then carefully examining them, without defining a quality benchmark. The standard cautions against the exclusive use of the screen for examination of a test run output.

The setup phase of Project Open Book utilized two discrete approaches to benchmarking as a way of transcending the limitations of the scanning quality standard. To establish the capabilities of the scanning system, the project team produced three test rolls of preservation microfilm, identical in all respects except for the average density readings. Reel 1 was exposed near the bottom end of the acceptable density range for preservation microfilm (.90); reel 2 was exposed near the top of the range (1.4); reel 3 was exposed in the mid-range of the average density guideline (1.1). A variety of standard test targets were filmed at the beginning and end of each reel, including AIIM X491 (Test target No. 1), AIIM X492 (Test target No. 2), and the RIT Process Ink Gamut Chart. Each reel also contained roughly 90 exposures from books, serials, archival records, and various illustrative materials representing the array of images likely to be encountered in Yale's existing collection of preservation microfilm. Following exposure, all three test reels were developed carefully according to the most stringent archival processing standards.[9] Duplicate negatives of each master negative reel served as the test rolls. The major advantage to the project in creating test reels was the possibility of controlling for the huge variation in the visual characteristics of original source materials, as well as controlling for subtle variations in the chemical processing of archival microfilm over time.

In order to establish the range of useful system image capture and enhancement pre-sets, the project team developed a second benchmarking system which will be used selectively in the production conversion phase of the project. In essence, the planned approach seeks to control for the capabilities of various display devices (CRT screens and laser printers). The fundamental assumption of this benchmarking process is that visual comparison of hard-copy output from the microfilm imaging system should be made neither against original printed source materials nor against a display screen, but rather against *digital images* of the original source materials. Furthermore, the approach assumes that 100 percent benchmarking of all converted materials is not necessary for establishing a viable quality control system. The following is a summary of the image benchmarking process planned but not fully implemented in the setup phase.

- Choose a sample of hard-copy originals, along with print negative counterparts.
- Digitize portions of the original volume at 600 dpi (title page, table of contents, selected illustrations, indexes) using a calibrated Xerox WG-40 flat-bed scanner with as many of the enhancement features invoked as possible and practical, and following the operational guidelines developed by Cornell University.
- Produce laser prints at 600 dpi on the Xerox DocuTech.[10]
- Digitize the identical pages from the microfilm print negative version.
- Produce laser prints at 600 dpi on the Xerox DocuTech.
- Compare matching prints under an eye-loupe (10X magnification), paying particular attention to letter fill-in or drop-out or fill in, highlights and shadows in line drawings, etc.
- Choose one combination of filter settings for the microfilm scanner that achieves most closely the appearance of the digitized original.
- Note the characteristics of the film source, once "maximum" quality has been obtained.
- Scan a volume with similar basic characteristics without benchmarking from the original.
- Compare prints of "benchmarked" volume with unbenchmarked one; adjust settings accordingly; note sources of discrepancies for future reference.

If the major goal of a conversion system is to maximize quality in a production environment, then it is essential to recognize that compromises will be necessary. The next phase of the project will begin to generate systematic information to answer the following questions: What is the nature of these compromises? Are they acceptable to users, to librarians, and to the preservation community? What are the advantages of the alternatives to the choices made for purposes of high-speed production? What research and development needs to be undertaken to establish parameters for machine-assisted quality control? Benchmarking from original source materials may be an acceptable solution in a pilot project, but in the real world of research, library preservation microfilm may be the ONLY form in which the original source now exists. We anticipate that the experience of converting 10,000 volumes from microfilm with only minimal reference to the original hard-copy published version will provide the library community with significant practical information on image conversion and image enhancement.

Digital Document Indexing

The project also seeks to address the question: "How can digital imaging technology provide enhanced intellectual access effectively and efficiently?" Intellectual access, in the context of Project Open Book, is defined simultaneously in two distinct senses. First, intellectual access is the suite of traditional practices of bibliographic control that organize books, serials, archives series, and other recorded knowledge in standard ways such as title, author, place and date of publication, and subject. When sophisticated database

management tools are linked to the capabilities of digital imaging systems, intellectual access has a second meaning, namely, the set of tools and techniques that permits users to grasp and make use of the content structure of the material, often represented through tables of contents, lists of illustrations or tables, and word indices to the full text of the material. In providing direct and ready access "inside" a published work transformed from the linear format of microfilm, digital imaging technology holds great promise of effecting the usefulness and useability of preserved materials. One challenge for Yale in Project Open Book is to identify efficient ways to use existing database technology and standards to develop structural indexes for large, complex, and varied published (and sometimes unpublished) source materials. More pointedly, the project needs to reconcile mid-nineteenth century publishing practices with late 20th century technology.

In its current single-platform configuration, the image conversion workstation provides for direct linkage to Yale Library's bibliographic database, Orbis, via a high-speed Ethernet connection. Library patrons will utilize the public access mode of Orbis to search among those portions of Yale's collections described online using the usual searching conventions, such as Boolean logic operators, to locate materials available in image format. The user interface, in its ideal configuration, would permit seamless hyper-linkages between the Orbis searching engine and the image server, allowing patrons to view image and index data for a title on a single access workstation without leaving the core program. The master plan for Project Open Book describes the way bibliographic, image, and index data would display in a distributed computing environment.[11] Currently, viewing of image files is only possible on dedicated access workstations, although the project is proceeding as if the overall goal will be met by the end of Phase IV.

In the setup phase, the production team established the mechanism for informing patrons of the existence of image formats and identified the procedures for establishing unique identification numbers for image files to provide the essential links between the image database and Orbis. Xerox's document management system is only used to log each item's unique identifier as it appears in the Orbis system and to log brief author and short title information. During the next phase, reformatting technicians will modify the bibliographic record for converted titles to indicate that an image version of the book or serial is available for consultation, using a standard phrase and noting the unique identifier of the image file.

Access to the content of an imaged document is provided by way of a separate relational database. Using Xerox's "Document Structure Editor" software, reformatting technicians apply page numbers, establish hierarchical nesting of images, and label the indexing levels as appropriate. In a volume scheduled for conversion, for example, the technician might identify the title page and associated prefatory materials, label the table of contents and other "finding aids" in the book, including back-of-the-book indexes and lists of illustrations, and indicate where each discrete chapter begins. A patron can then go

directly to any of these structures.

The structural information is contained in an RDO (Raster Document Object) file, which is Xerox's adaption of the proposed international Office Document Architecture (ODA) and Interchange Format. An RDO file contains information about the structure of an image document, as well as a file location pointer for each page image in that document. Each page in the document is represented by a single TIFF (Tagged Image Format File) file. A TIFF file contains the digital data from the scanned page, along with a TIFF header which describes the characteristics of the image file, including the formula used to compress the digital data. A TIFF header is comparable in scope and function to the fixed field header in a MARC bibliographic record, and is analogous to an envelope that addresses and describes the contents in a structured way. Because the Xerox database system is partially proprietary, however, the structure of a complex image document, such as a book, is displayable to the patron only with the use of Xerox software.

During the setup phase, the project team experimented with a number of approaches toward applying the document structure editor to a wide variety of mid-nineteenth and early twentieth century published materials. This was a period during which authors and publishers themselves experimented with the structure of the books they produced. Tables of contents varied widely in their placement and level of detail. Similarly, word indexes to the content of a particular work took many forms and served a variety of purposes not commonly seen today. Special lists of illustrations, glossaries of terms, and other internal tools aided the reader in using a volume. As the next phase of the project begins, Yale's production team will seek ways to apply the structural editor consistently, while recognizing the intense inconsistency of the forms and formats of the materials being converted. At the very minimum, the actual page numbers of the volume will be applied and the images for a particular volume will be clustered under the title page, which will display first to the reader when the document is retrieved. If resources permit, at least two intermediate "levels" of indexing will be applied. The first of these will involve identifying and tagging each volume's internal "finding aids," including tables of contents, word indexes, and special lists. A second level will cluster pages in logical intellectual groups, such as chapters, sub-parts, articles, or issues, based upon the specific structure of the work. Our report on the next phase of the project will provide more detail on indexing structures utilized.

Conversion Workflow Management

The operational core of Project Open Book is sustained production-conversion of microfilm frames to digital images. The central workflow question of the setup phase was "How might quality image conversion and indexing be obtained in a production environment?" The project team isolated and investigated a number of independent factors that have a major impact on the rate of conversion. Among the most important of

these factors are the technical limitations of the scanning hardware (particularly the mechanics of winding and re-winding film on the scanner); incomplete bibliographic records and technical targets that hinder the preparation stage; the huge variation in the technical characteristics of preservation microfilm; the complex structure of many published materials or the nearly complete absence of structure (including page numbers) for others; weaknesses in the user interface that complicate quality assurance activities (especially the complexities of the zoom feature and the control of default filter settings); and, slow data transmission speeds on the network and within the integrated software system. During the next phase of the project, we anticipate that the networked, multi-workstation environment will produce an additional set of project management issues that will be described in subsequent reports.

During the setup phase, the project team outlined in some detail the four stages of the conversion process and estimated the productivity rates of each of these stages. The conversion of 35 millimeter microfilm frames to digital images in some ways mirrors the initial creation of preservation microfilm itself. The **Preparation Stage** begins with the application of the content selection criteria and ends with the mounting of the film on the conversion hardware equipment. The key preparation steps are:

- producing a print-out of the appropriate bibliographic record for the title;
- marking the location of the first and final frames of each volume on the reel with removable silver tape;
- cleaning the film of dust; and,
- inspecting the microfilm on a light-table and noting a range of characteristics on the workform.

Among the most important characteristics of the filmed item that are noted on the workform are: reduction ratio, orientation, contrast, print type and size, condition of original, and the frequency of illustrations. Five density readings are taken for each volume and noted on the form, along with the average of these five readings. Additionally, any special instructions for the reformatting technicians are noted on the workform, including obviously missing pages (often indicated by special targets inserted in the film sequence), gaps in the filming, excessive splices, and the actual page number of the last frame of the volume. Finally, a simple mock-up of the first and last pair of pages for the item is sketched on the workform. The silver tape and the mock-up are required because the Mekel scanner is not equipped with a mechanism for viewing film that is mounted on the scanner.

Once the reel of microfilm has been mounted on the calibrated scanner, the **Scanning Stage** can begin. The key steps in the scanning process include creating a batch directory for the item with a unique file name, establishing the scanning cropping parameters based upon a test scan, selecting scanning options by entering the results of the film inspection

from the workform, and beginning scanning in a continuous mode. At its maximum capture rate of 600 dots-per-inch, the Mekel scanner actually converts one-half of each frame, representing an individual page, in a single pass through the length of the film. At the end of the volume, the technician rewinds the film to the beginning of the volume and resets the scanner so that the second half of the frame, again representing a single page, can be scanned. This technical solution for maximizing the resolution of digital conversion requires that materials on the microfilm be oriented in a "two-up" cine mode, as opposed to a comic mode. The "two-up" orientation allows the scanner and associated software to separate two pages filmed as a single microfilm frame. Even though the conversion software provides for the scanning of individual page images in comic mode, the Mekel scanner can obtain a maximum resolution of only 300 dpi.

During the scanning stage, the reformatting technician may re-scan an unacceptable image at any time and overlay the newly converted image on the old. The major reasons that re-scanning may be necessary include the failure of the edge-detection software, the presence of a particularly complex illustration, a suddenly poor alignment of the original item on the film, or skewing of the original in excess of 10 degrees. The scanned "batch" must be imported into the Xerox Documents on Demand software system from the TurboScan software associated with the microfilm scanner. The Xerox software converts the batch of discrete image files into a single image document (called a Raster Document Object or RDO) that can be managed as a single object.

Once the scanned batch has been successfully imported, the **Indexing Stage** may begin. Yale's report on the next phase of Project Open Book will included a more complete description of the indexing process. During the indexing stage, the technician has the best opportunity to undertake image-by-image quality control, including the generation of proof-prints of selected images. Unacceptable images can be re-scanned immediately or flagged for re-scanning at a later stage of the process. The setup phase demonstrated that, given the technical configuration of the system, the most cost effective quality control takes place at the point of conversion, rather than farther down the production line.

The final steps in the production-conversion process are those of the **Quality Acceptance Stage**, in which the quality and completeness of the imaged and indexed document is confirmed and the data saved to optical disk. The major steps in the stage include proof-reading the accuracy of the structural index information and the pagination of the image document, producing selective benchmark printouts to re-check the quality of the image conversion process, re-scanning of highly problematical images, as necessary, transferring image and index data to magneto-optical disk, and the registering of the completed document in the Xerox Document Manager software. During the registration step, the technician enters the key descriptive attributes of the completed image document that will be required for searching and retrieval by users until seamless linkages are made with the Library's online catalog.

During the setup phase, the project team developed and evaluated two possible workflow models and validated an equipment configuration that supported the team's estimates of the best workflow model. The team also compiled statistics on the rate of the overall conversion process, concentrating especially on the time required to convert microfilm frames in "continuous" mode. Based upon the experience of converting 100 volumes in a production-conversion mode, we estimate that 60 minutes will be required to complete the conversion of a typical 300-page volume from microfilm to digital imagery. This hour-long process breaks down as follows: 8 minutes for preparation activities, including visual inspection, density readings, workform completion, and preset selection; 25 minutes for digital conversion; 15 minutes for content indexing; and 12 minutes for quality assurance. Excluded from these estimates are a variety of document management activities that currently are quite time consuming and, therefore, will most likely be carried out in batch mode after hours. These tasks include file transfer from magnetic to optical disk and file backup.

Selection for Conversion

The setup phase defined the components of an issue that is critical to the overall success of Project Open Book, namely, a selection process grounded in specific selection criteria. One of the four working hypotheses of the project is that scholars will find greater benefits from a cohesive, concentrated, and significant body of digital information. In the abstract, the question "What is the virtual library?" is far too complex a matter to be addressed meaningfully in a single pilot project at a single research library. A guiding principle of Project Open Book, therefore, is that the selection issue should be focused, in operational terms, on an existing corpus of preserved materials whose long-term value has already been determined, and then further defined by identifying cohesive subsets of this corpus that can meet present teaching and scholarship needs of the university. Finally, the field of materials qualified for conversion will be narrowed by applying technical criteria designed to maximize the quality of the converted microfilm image and the productivity of the conversion process.

Existing Microfilm: Yale will not create preservation microfilm specifically for Project Open Book. The selection process is simplified for the project, therefore, because Yale will draw on the rich collection of microfilmed monographs and serials that it has created over a decade of large-scale brittle book preservation projects funded in part by the National Endowment for the Humanities. In general these collections include items from the period 1830 to 1950 that have already had significant curatorial review and been judged worthy of long-term preservation. Among the microfilmed collections that will be tapped for digital conversion are the 25,000 volume American History monographs collection, the 22,000 volume European History collection (excluding the United Kingdom), the 3,200 volume French History collection, and the 10,000 volume History of Economics and Political Science collection. This latter collection is part of an ongoing

NEH project, and another 10,000 volumes will be filmed in the next year. In addition to these large and intellectually cohesive collections, Project Open Book has at its disposal some 5,000 microfilmed volumes of high-use items that were preserved on demand from the entire corpus of humanities collections at Yale.

Content Usefulness: First and foremost, Project Open Book is a pilot research and development program within the Yale community. As such, ongoing evaluation by a subset of potential users of the image library is crucial to the success of the project. The image library created in the project must be used relatively frequently in order for a systematic evaluation program to be implemented. Content usefulness as a selection criterion, therefore, is defined as those clusters of materials from the overall microfilm pool with known relevance to the ongoing teaching and research program of Yale's humanities scholars. Furthermore, the relevance of the collection must be tied to a commitment from university scholars and students to participate aggressively in the evaluation program. Selection by content usefulness will not involve item-by-item decision making, but instead will be a matter of identifying classification clusters, from the Old Yale and Library of Congress schemes, within large pools of preservation microfilm that hold promise of contributing to ongoing teaching and scholarship over the next two years.

Technical Characteristics: Given the known technical limitations of the hardware/software configuration, a set of basic technical selection criteria must be applied to the microfilm collection, on a title-by-title or reel-by-reel basis, prior to acceptance for conversion. Selection by technical characteristics of the input source allows for the convergence of maximum conversion quality at an optimum production rate. The overall goal of the selection process is to be as inclusive as possible and to attempt to respond to any special technical challenges that arise during the course of routine conversion. Any given volume will not be excluded capriciously for technical reasons. Among the basic technical criteria that will be applied to each reel and/or title that has passed through the "content" filter are:

- *Film Stock*: Input source will be 35 millimeter silver halide microfilm. The duplicate negative (2N) should be used for conversion. Conversion from positive microfilm (3P) is slower and of slightly lower quality because of image degradation in the copying process. The master negative (1N) should never be used for digital conversion, given the risk of damage to the preservation copy.

- *Image Density*: The average density readings from the technical targets should be within the range specified in the *RLG Preservation Microfilming Handbook* for normal exposures (1.0 to 1.30). A minimum of two readings should be taken per title from the technical target, with a preference for five readings from two separate technical targets for each title. Additional readings from the body of the item itself

should be obtained as needed to construct a full portrait of the volume's density.

- *Orientation:* Original materials should be filmed "two-up" in cine-mode with a limited number of full-frame exposures from any particular volume. Random full-frame exposures (e.g., fold-out maps or illustrations) require interrupting the continuous scanning mode of the equipment, manually adjusting the settings, and then re-setting the software to resume continuous mode.

- *Reduction Ratio:* An accurate measure of the reduction ratio is crucial to the proper conversion of microfilmed materials. The reduction ratio may either be obtained from the technical targets that precede the filmed item or from the bibliographic record. In the absence of the reduction ratio figure noted at the time of filming, the reduction ratio may be calculated by reference to the original item. The time to re-calculate the reduction ratio in this manner, however, may add substantially to the cost of converting the item.

- *Condition of Film:* Film stock should be relatively clean and free of scratches, redox blemishing, or major processing water-spots. Additionally, because film splices may result in scanner error, the number of internal splices (as opposed to splices between volumes) should not exceed RLG's recommendation of six for the entire reel.

Additionally, the setup phase has demonstrated that the quality and rate of digital image conversion are substantially improved when the following characteristics are present in the microfilm source:

- *Consistency of average density* across the width of the reel, across the length of the title, and from title-to-title throughout the reel. Image conversion proceeds more rapidly and more reliably the greater the consistency of microfilm image quality across the reel.

- *Clearly defined edges* in the original source material filmed, as measured by a rapid shift from black to white on the film. The accuracy of page-by-page conversion depends in part on the implementation of edge-detection software. If original materials were severely yellowed (faded or discolored) or if the edges were cracked, chipped, "dog-eared," or otherwise not clearly defined, the edge-detection software can be "fooled."

- Wide *exposures* complicate the system setup procedures and, therefore, increase the time required to convert the volume. Frame exposures should not be set nearer than three millimeters from the edge of the film.

- The *gutter* of a bound volume should be rigorously centered in the film aperture.

Avoid "creep" of the center line of the material across the film, which slows production as scanner settings are readjusted.

- The *alignment* of materials should be "square" to the edge of the film. Although software can "de-skew" a digital image tilted up to 10 degrees from square, the process is very time consuming and therefore expensive.

Although the absence of one or all of these factors should not necessarily be grounds for exclusion from conversion, it is important to recognize that the cost of working around the limitations implied by the absence of a consistent, high-quality input source (whether microfilm or paper) most likely will either substantially increase the cost of conversion, due to the necessity to devote extra time and effort to "tweaking" individual images, or will substantially lower the overall quality of the converted images.

Project Evaluation

Given the fundamental principle that the digital library created in Project Open Book must have value to contemporary users in the Yale Library, we expect the selection of materials for conversion to proceed in tandem with the process of project evaluation. The focus of the evaluation is on the useability of the system in enhancing access to preserved materials and the usefulness of the materials themselves for scholarship when converted to image form. Our principal hypotheses is that digital image systems are at least as useful for research and at least as useable as their hard copy and microfiim counterparts for researching the same topic. The clusters of issues to be investigated and systematic information to be gathered include the usefulness for research purposes of image content, document structure indexes, and linkages to the local online catalog, compared to traditional access and browsing methods; the usefulness and quality of the user interface, screen displays, and local printing capability, as well as the need for these and additional tools and capabilities; success in retrieving relevant information sources when the structure and overall content of the database is largely known; and, the characteristics of the user population, especially its visual grasp of display screen layout, familiarity with content of the image database, and range of experience with online searching or the use of research resources in electronic form.

The project team, with the help of the Steering Committee, has begun to identify user groups in the Yale community who are willing to test the usefulness of the materials selected for conversion and the useability of the end product of the conversion process itself, including the document structure editing, and the user interface. One such user group will consist of a faculty member and students enrolled in specific courses involving primary research in the humanities. Material relevant to the class would have been digitized earlier in the project and the use of the digital library would be a required assignment for at least half of the class. The second half of the class would be required to

use materials on microfilm and/or original bound formats. Class members would provide feedback on the nature and quality of the imaging system and the research papers produced in the course could be evaluated for their use of image-based, as well as traditional library materials. Such modified "controlled field experiments" promise to provide valid and reliable assessments of both the useability of the system and the usefulness of the materials in the digital library for scholarly research.

Conclusion

Yale University Library is now prepared to embark on the third phase of Project Open Book--the production-conversion of 3,000 volumes of the projected 10,000 volume digital library. In this phase Yale will identify the requirements for effective and economical digital conversion from microfilm images. Given the optimal scanner settings and the workflow configuration as determined during the setup phase, work in the third phase will validate the production-conversion model developed in the setup phase, attempt to resolve any problems that arise from a three-workstation environment, and seek further to optimize productivity in the process.

The third phase will also examine the requirements for intellectual access to the digital library through the addition and use of electronic links between the scanned images and the standard apparatus--pages, tables of contents, indices, chapters, and so on--by which complex documents are typically organized. This phase will systematically incorporate user access capabilities into the project. Yale will add a view/access workstation in the library during this phase and will also introduce a network accessible document server with an optical disk jukebox. The introduction of wider access to the digital library will enable Yale to execute its plans to evaluate user responses to the imaging system during this phase. Yale must be able to understand and demonstrate its ability to provide the image document library as a secure and dependable network service. During the next phase, Yale will begin to develop such an ability.

The growing document library and the enhanced conversion, storage, and access subsystems achieved in the third phase of Project Open Book will create a solid foundation for multiple and concurrent networked access to the image library. In the final phase of the project--the networked distribution phase--Yale will continue the high-volume effort begun earlier and provide access to the growing digital library over the campus network. Formal evaluation of user responses will continue. Yale will explore issues of network access and consider the costs and benefits of using service bureaus for digital conversion and document structure editing.

The deliberate phased approach of the project is a pivotal component of the overall strategy to achieve the critical objectives of the project. At the completion of Project Open Book in 1996, the Yale Library, with its vendor and university partners and with the

support of its collaborators at other institutions, will have thoroughly examined the means, costs, and benefits of converting large quantities of preserved library materials from microfilm to digital imagery.

Acknowledgements: We wish to thank the Commission on Preservation and Access for its support of Project Open Book. The support staff of the Xerox Corporation, in particular Rob Martella, were responsive to our needs for training, technical assistance, and product upgrades. Similarly, the Amitech Corporation provided ongoing technical support for the microfilm scanning hardware and software. At Yale, we gratefully acknowledge the support of the members of the Project Open Book Steering Committee. Finally, the setup phase was in large part made possible through the willing and enthusiastic work of the Project Team, composed of the authors of this report plus Donald J. Waters, Greg Kaisen, and Robert Halloran.

Project Open Book Conversion Workstation

- *Capture Station*

  Hardware:

  AST PC 486/33, 16 MB memory, 1.05 GB hard drive, 3.5" 1.44 MB floppy disk drive, 5.25" 1.2 MB floppy disk drive
  Cornerstone Dual Page 120 19" monochrome monitor
  Serial mouse
  Netflex Ethernet interface card
  Xionics Turbo Graphics accelerator board
  UltraStor SCSI controller board
  Xerox 4030 II laser printer

  Software:

  Microsoft DOS ver. 5.0
  Microsoft Windows ver. 3.1
  Xerox Documents on Demand software
  Xerox Postscript Integration System software
  Gupta Technologies, Inc. SQLBase ver. 5.0
  Beame & Whiteside TCP/IP communications software ver. 3.0

- *Microfilm Conversion Sub-system*

  Hardware:

  Mekel Engineering 400XL Microfilm Digitizer
  Hewlett Packard LaserJet III

  Software:

  Amitech Turbo Scan ver. 3.0
  IPT Scan Optimizer ver. 6.0

- *Optical Storage Sub-system*

  Sony 5.25" Optical Drive
  Sony 5.25" Magneto Optical Disks (EDM-1DA0s)

The following technical manuals document the operation of the hardware and software components of Yale's digital imaging system.

- *Operating Instructions and Maintenance Manual for the Mekel M400XL Microfilm Digitizer*, Manual #5052, 1989. (Mekel Engineering, Inc., Diamond Bar, CA)
- *Amitech Turbo Scan Software, V3.0.* Installation and Operation Manual. (Amitech Corporation, Fairfax, VA)
- *IPT Scan Optimizer, V6.0., 1991.* (Image Processing Technologies, Inc., Vienna, VA)
- *Xerox Documents on Demand, V2.0C*, 1994. XDM/XQM User Guide. (Xerox Corporation, Webster, NY)

Endnotes

1. Donald J. Waters, *From Microfilm to Digital Imagery: On the feasibility of a project to study the means, costs and benefits of converting large quantities of preserved library materials from microfilm to digital images,* (Washington, D.C.: Commission on Preservation and Access, 1991).

2. Anne R. Kenney and Lynn K. Personius, *A Testbed for Advancing the Role of Digital Technologies for Library Preservation and Access: Final Report by Cornell University to the Commission on Preservation and Access,* (Washington, D.C.: Commission on Preservation and Access, 1993).

3. Donald J. Waters and Shari Weaver, *The Organizational Phase of Project Open Book,* (Washington, D.C.: Commission on Preservation and Access, 1992).

4. *Glossary of Imaging Technology,* Technical Report TR2-1992, (Washington, D.C.: Association for Information and Image Management, 1992).

5. *Recommended Practice for Quality Control of Image Scanners,* ANSI/AIIM Standard MS44-1988, (Washington, D.C.: Association for Information and Image Management, 1988).

6. Nancy E. Elkington, editor, *RLG Preservation Microfilming Handbook,* (Mountain View, CA: The Research Libraries Group, 1992).

7. *Resolution as it Relates to Photographic and Electronic Imaging,* Technical Report TR26-1993, (Washington, D.C.: Association for Information and Image Management, 1993).

8. *Standard Recommended Practice -- Monitoring Image Quality of Roll Microfilm and Microfiche Scanners,* ANSI/AIIM Standard MS49-1993, (Washington, D.C.: Association for Information and Image Management, 1993).

9. *Practice for Operational Procedures/Inspection and Quality Control of First-generation, Silver Microfilm of Documents,* ANSI/AIIM Standard MS23-1991, (Washington, D.C.: Association for Information and Image Management, 1991).

10. Patricia J. Smith, "Xerox's DocuTech: 'Print Shop' All in One," *The Seybold Report on Publishing Systems* 22 (22 June 1993): 3-21.

11. Waters, *From Microfilm to Digital Imagery,* pp. 25-6.

For more information on The Commission on
Preservation and Access and its publications
Call or write -
The Commission on Preservation and Access
1400 16th Street, N.W., Suite 740
Washington, DC 20036-2217
(202) 939-3400